ENERGY STAR Connected Thermostats

Stakeholder Working Meeting

January 23, 2020

Attendees

Abigail Daken, EPA
Abhishek Jathar, ICF for EPA
Alan Meier, LBNL
Leo Rainer, LBNL
Eric Floehr, Intellovations
Craig Maloney, Intellovations
Michael Blasnik, Google/Nest
Jing Li, Carrier
Brian Rigg, JCI
Theresa Gillette, JCI
Rohit Udavant, JCI
Kurt Mease, JCI (Lux)
Diane Jakobs, Rheem
Carson Burrus, Rheem
Chris Puranen, Rheem
Glen Okita, EcoFactor
Brent Huchuk, ecobee
John Sartain, Emerson
James Jackson, Emerson
Mike Lubliner, Wash State U
Charles Kim, SCE
Michael Fournier, Hydro Quebec
Dan Fredman, VEIC
Robert Weber, BPA
Phillip Kelsven, BPA
Casey Klock, AprilAire
Wade Ferkey, AprilAire
Ulysses Grundler, IRCO/Trane
Jeff Stewart, IRCO/Trane
Mike Caneja, Bosch
Sarathy Palaykar, Bosch
Brenda Ryan, UL
Mike Clapper, UL
Alex Boesenberg, NEMA
Ethan Goldman
Jon Koliner, Apex Analytics
Michael Siemann, Resideo
Aniruddh Roy, Goodman/Daikin

Agenda

- Recap of RHU decisions and questions
- Update on process for controllers of variable speed HVAC
- Software
  - Version 2 implementation (discussion we ran out of time for last month)
  - Documentation: Inform, ask for feedback.
  - Test Data Set
- Ethan Goldman: CalTRACK hourly methods and potential relevance for ENERGY STAR metric improvements (Did not get to this topic, slides removed)
- Other topics as time allows

RHU: Where were we again?

- RHU covers a perverse incentive in the metric: deeper longer setbacks with ER heat for recovery will show up with higher % heating run time reduction, but actually use more energy

- Resistance heat utilization (RHU) is a measure of how well thermostats avoid use of resistance heat backup for heat pumps

$$RHU = \frac{resistance\ heat\ run\ time}{total\ heating\ run\ time}$$

- Calculated for homes with a heat pump and resistance back up (per tstat wiring); calculated daily, average over year w/in outdoor temp bin (e.g. 30F - 35F)
- RHU2: Two improvements exclude some installations from the average
  - < 30 total heating runtime hours in bin (conditions heating system is not designed for)
  - Run time in top 5th percentile w/in bin (malfunctioning heat pump)

RHU2: Recap

- Previous meetings: shared RHU2 results, discussed need for additional sampling of heat pumps for reasonable results
- Metric proposal: RHU2 in 30F-45F bin, upper 95th confidence limit of the mean ≤ 0.2
  - Question: National vs. regional roll up: More samples in the mixed humid region contributes to low upper 95th confidence mean values
  - Question: 5F wide or 15F wide bin(s)

    Generally a gradual increase in the number of samples with temperature

    Installations from different vendors in each bin are similarly distributed across regions

    Large samples help because assuming overall vendors have installations that are in the same population, larger samples will be a more accurate representation
  - Question: Criteria on a Percentile vs on the Mean (or it's confidence bound)

    Percentiles less influenced by extreme values, but throw out info in lower percentiles

    Mean keeps info, and RHU2 already trims installations w/ top 5% of RHU before calculating the mean

RHU2: Recap

  - Question: weight RHU with something? If there's 0% ER off time in the day, the ER minutes don't count (broken system); if there was heat running only 50% of the time, then you'd get half weighting. If heat running very little, it would be heavily weighted.
- Previously discussed outlier considerations
  - User settings that could cause that to happen? e.g. compressor lockout temp, the way you set the flipped behavior, etc.
  - Comfort setting may also effect this – which contractors would like.
  - Use an outlier threshold – if lots of compressor runtime, no need to include it in the average.

- Asking for heat pump only sample with February resubmission – we haven't done this with RHU2, and hope it will help answer national vs. regional question and 5F vs. 15F bins.

Compressor and Emergency Heat Runtime Variables

- There are three runtimes used for the RHU calculation defined in the *Version 1.0 Method to Demonstrate Field Savings of ENERGY STAR Connected Thermostats*:
  - **t_{comp}** = total compressor heating run time (a proportion of t_{comp} may occur when auxiliary resistance heating also occurs)
  - **t_{aux}** = total auxiliary resistance heating run time. Use of electric resistance strip heat to increase heating capacity.
  - **t_{emerg}** = total emergency resistance heating run time. Use of electric resistance strip heat under fault conditions or when the outside temperature is very low.
- There are three runtimes in the thermostat interval files which are defined in the *Thermostat Interval Data* section of the software documentation:
  - **heat_runtime** - daily runtime (in minutes) of the heating equipment.[3]
  - **auxiliary_heat_HH** - hourly runtime (in minutes) of auxiliary heat equipment (HH = 00-23).
  - **emergency_heat_HH** - hourly runtime (in minutes) of emergency heat equipment (HH =

[3] Should not include runtime for auxiliary or emergency heat - this should be provided separately in the columns *emergency_heat_HH* and *auxiliary_heat_HH*.

Compressor and Emergency Heat Runtime Variables (2)

- The **auxiliary_heat_HH** and **emergency_heat_HH** inputs are unambiguous and map directly to $\mathbf{t}_{aux}$ and $\mathbf{t}_{emerg}$ respectively.
- The **heat_runtime** input could be interpreted as either:
  - "total compressor time" (the same as $\mathbf{t}_{comp}$)

  or
  - "compressor only time" ($\mathbf{t}_{comp}$ - $\mathbf{t}_{aux}$).

- Questions:
  - Which do you store in **heat_runtime**?
  - Which do you think makes the most sense (in terms of calculating RHU and the Metric)?
    - In RHU calc, just need to make sure the instructions and the code match
    - In heating % run time reduction metric, the heating run time is used in the fit correlating the temperature demand to the equipment run time

Discussion: RHU

- Question on RHU vs. RHU2: do we imagine one metric in the end? Yes, no idea what it would be called.
- Any vendors know whether they are interpreting t_heat as total compressor time or as compressor only time?
  - 1 vendor is using total compressor time (t_{comp})
- Which makes sense for the metric calculation?
  - Including the compressor + resistance heat run time in the fit will introduce large variations in capacity from data point to data point, BUT
  - Excluding aux run time may also distort the fit, because as the indoor-outdoor temperature difference goes up, anomalously short run times would seem to be used
  - Could confine the fit to days or hours when only compressor was used, at the cost of excluding data
  - Could back out to whatever we need for the fit one way or another
  - Matching however vendors store their data minimizes opportunity for errort

Controllers for Variable Speed HVAC

- Reminder: Current idea is to treat communicating controls separately
  - Field data analysis to show significant time in lower capacity states
  - Controls verification procedure to confirm that (fixed) setting used in HVAC efficiency test are achieved by native control
  - Potential allowance for additional standby power; discussion of other spec criteria
- Met with AHRI Unitary Regulatory Committee 1/7/20 to discuss concept outlined above
  - No fundamental objections raised
  - Agreed to work together
- More from AHRI members on CVP
  - Original intent: confirm that the HVAC ratings will be achieved in the field, paradigm shift to use for evaluating controls
  - Designed for commercial equipment and would need modification for residential
  - Needed when field data analysis may serve similar purpose?

Controllers for Variable Speed HVAC

Workshop March 26 9:00am to 3:00pm
ASHRAE Headquarters, Atlanta
All are invited, look for RSVP in a few weeks
Remote participation will be available

- Likely pre-work:
  - EPA may suggest a metric (or several) as a straw man; if so, come ready to discuss
  - AHRI members to review CT V1 specification and note other criteria that may be a barrier to participation
  - AHRI will discuss modification of CVP
- Breaking news (this morning): Draft 2 version 6 ENERGY STAR CAC/HP spec proposes low ambient CVP for residential cold climate heat pumps to be developed by DOE

Discussion: Communicating controllers

- Clarify: not a separate specification, but potentially completely separate metrics and performance requirements within the same spec

Software: 2 Stage systems implementation

- Seeking to include installations wired to control 2-stage heating and maybe cooling.
- Furnace/boiler: efficiency independent of capacity, no issue. Compressor based heat/cool: difference in efficiency will cause systematic nonlinearity. Include these?
- Vendor feedback from previous meetings: too few samples to bother?
  - Low single digit installations wired for control of 2 stage heating (DIY installations)
  - Another vendor: 15% for model that is contractor installed, 7% for DIY
- Proposed implementation: change input file format to have extra columns
  - Heating run time -> equivalent full load run time (ERT)
  - Add columns for stage 1 and stage 2 heating run time
  - If ERT column filled out, use that (assume calculated by vendor using actual relative capacity of stages). Document expectation of how it's calculated?
  - If no ERT data, calculate from stage 1 and stage 2 run times using 0.65 relative capacity
  - Also add installation wiring types for 2-stage furnace/boiler heating

Discussion from December meeting: 2 stage implementation

- Vendor experimented with various ways of doing this, and found that it didn't make much difference to the scores. The choices of relative capacity affect the slope, but since that drops out of the performance metric, it just doesn't make much difference. Argument for just not worrying about it much. Difference between comfort temp and average temp drives savings, with a consideration of climate.
- In cooling analysis had a slightly bigger impact, but still very small (couple of tenths of % difference on score)
- Any further discussion? (Cut off short)

Further Discussion: 2 stage implementation

- Even for cooling systems, efficiency may not change with capacity, based on how you design the system. Furthermore for fossil fuel heating system, you can get variations in efficiency if you can't change the inducer fan speed in time with the fuel input
- Are we talking about input capacity or output capacity?
  - Output capacity may be better correlated with load; run time is correlated with output capacity, so using input capacity includes additional opportunities for savings that we haven't otherwise explored
  - Input capacity is what we really want to know, because that's where the savings are
  - What is for vendors to get and analyze? Nominal input capacity. Natural gas heat content isn't the same every place, and installers may not adjust it properly, but that wouldn't effect relative capacity. Compressor based systems you'd know nominal relative output capacity. Neither nominal capacity takes into account that efficiency varies from one time to another.

Further Discussion: 2 stage implementation

- What is our goal? Are we trying to capture efficiency from using lower stages, or just better calculate the schedule efficiency?
  - Multiple stages does afford some opportunity for savings, through avoiding cycling losses
  - Thermostat vendors have in the past have talked about control strategy to run the fan a little longer to harvest energy as the heat exchanger is cooling down/heating up
  - For some systems, there are efficiency opportunities in the lower stages… aren't there? Fan energy is complicating too – lower speed taps on fans are less efficient. Also energy loss in the ducts may be affected by fan speed and it isn't even clear what the affect is.
  - As the metric stands today, we don't have a baseline of comparison the a single speed systems.
  - EPA's goal is the better calculate the schedule efficiency for these systems

Further Discussion: 2 stage implementation

- Whether we set a criterion or not, is it worthwhile to track ERT/RT or some other simple metric?
  - Unfortunately, that **is** sensitive to whether you get the relative capacity right
  - Could use a ratio of low and high stage run time – not sensitive to our assumption about relative capacity, but definitely sensitive to the **actual** relative capacity.
  - .65 much more accurate for furnaces
  - Setback also has an effect on this, as high stage may be used to recover from set back
  - AFUE calculation assumes that the furnaces (and boilers) runs in low fire all the time (or the vast majority of the time), so time in low stage is a measure of whether consumers are getting the efficiency they expect
  - Boilers, even two stage, are often wired as single stage (confirming with other vendors)

Software: Version 2.0 implementation

- Reminder: we talked about changing input format to be all hourly data, no daily, and this was generally agreed to be a good plan
- Hourly data will change calculations somewhat (see next discussion point)
- While we want to test as many V2.0 spec changes as possible with resubmission data, this may be too much to fit within the version 1.x software thread
- As a starting place for discussion EPA proposes:
  - Start Software version 2.0 branch now; not to be used for certification or resubmission data. Use it only for testing
  - Summer 2020 resubmission to include results from both version 1.7.0 and 2.0?
  - Heat pump oversample run, as proposed for V2 Method, used to calculate RHU
  - Development efforts kick off in Q1 2020

Software: Calculating temperature based on hourly timeline

- What do we take in hourly vs daily?
  - Hourly data : Average indoor and outdoor temperature, Auxiliary heat run time, Emergency heat run time, Average conditioned space temperature, Average heating/ cooling set point temperature.
  - Daily data : Heating/ Cooling equipment run time
- Effect on calculation of final metric?
- How would it look different if we used hourly data?
  - ER use is compared to run time in a day
  - Possible to track recovery?
  - Hourly baseline?

Discussion: V2.0 software implementation plan, and hourly data/calcs

- No comments at this time

Software Documentation: Plan for comment

- Method documentation will be maintained by EPA
  - Available from the EPA partner site: https://www.energystar.gov/products/connected_thermostats/partners
  - Updated with each major version update
- Read the Docs (directly pulled from documentation in the code) will be maintained by Intellovations
  - https://thermostat.readthedocs
  - From the code in the GitHub repository: https://github.com/EPAENERGYSTAR/epathermostat
  - Updated with each minor version update

- Discussed difficulty of keeping separate Read the Docs documentation and the more narrative explanation in the Method, but ultimately felt both had their place
- Any feedback?

Thermostat Test Data Set

- Current
  - Consists of synthetic test data
  - Used for flagging changes caused by code updates (regression testing)
- Proposed new set
  - Composed of anonymized samples of real thermostat data
  - Investigate edge cases
- Range / Diversity of:
  - Meta data (vendors, product types, climates, HVAC capacities, HVAC types)
  - Inputs (set points, temperatures)
  - Outputs (metric, RHU, model slope)
  - Operation (cycling, schedules, equipment failures)

Discussion: Test data set

- Goals:
  - Traditional regression testing: we changed the software and expected the results not to change or to change in a known direction, so we can see whether that happened
  - From several vendors, to make sure we're not being accidentally unfair
  - Thermostats that tend to expose bugs in the code, to make sure that those bugs do not recur in new versions
  - Right now, it's a long cycle to test data at all, and ends up with no insight into effect on individual thermostats
  - Also may help us understand the effect of filtering, protect against problems with the metrics, etc.
- This would be in the form that it goes into the module

Discussion: Test data set

- Would vendors be willing to contribute to this data set? Or, what protections would you need to feel comforting? If you've shared data before, what kind of legal or administrative steps were needed.
  - Vendors haven't checked in with the legal department, but might be interested in helping.
- Clarification: actually need LESS metadata than is in Donate Your Data set; otherwise similar
- For any vendors that have shared data with researchers, what arrangements did you make?
- Potential for thermostat vendors to pass on a request from EPA?

Discussion: Test data set

- For vendors' internal research, they may use data sets from friends and family early adopters, who have different expectations and agreements about privacy.
  - That would be fine.
- There may be data sets that have been shared with a utility, where customers also have different expectations of privacy.
  - That would be fine too.
- In fact, having different sources from different vendors is also not a problem.

Wrap up and Next Steps

- Action Items:

- Next Steps: